0590
0424

3

OIPE

ENTERED

RAW SEQUENCE LISTING DATE: 04/25/2002 P.6
PATENT APPLICATION: US/10/054,841 TIME: 12:04:27

Input Set : A:\19999YCA.txt
Output Set: N:\CRF3\04252002\J054841.raw

<110> APPLICANT: Chen, Fang
<120> TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR RECEPTOR PROTEINS
<130> FILE REFERENCE: 19999YCA
<140> CURRENT APPLICATION NUMBER: 10/054,841
<141> CURRENT FILING DATE: 2002-01-23
<150> PRIOR APPLICATION NUMBER: 09/487,379
<151> PRIOR FILING DATE: 2000-01-18
<150> PRIOR APPLICATION NUMBER: 09/141,000
<151> PRIOR FILING DATE: 1998-08-26
<150> PRIOR APPLICATION NUMBER: 60/078,633
<151> PRIOR FILING DATE: 1998-03-19
<150> PRIOR APPLICATION NUMBER: 60/062,902
<151> PRIOR FILING DATE: 1997-10-21
<150> PRIOR APPLICATION NUMBER: 60/057,090
<151> PRIOR FILING DATE: 1997-08-27
<160> NUMBER OF SEQ ID NOS: 30
<170> SOFTWARE: FastSEQ for Windows Version 4.0
<210> SEQ ID NO: 1
<211> LENGTH: 2807
<212> TYPE: DNA
<213> ORGANISM: Human
<400> SEQUENCE: 1

gaatatgatg accctaatgc aacaatatct aacatactat ccgagcttcg gtcatttgga 60
agaactgcag attttcctcc ttcaaaatta aagtcaggtt atggagaaca tgtatgctat 120
gttcttgatt gcttcgctga agaagcattg aaatatattg gtttcacctg gaaaaggcca 180
atatacccag tagaagaatt agaagaagaa agcgttgcag aagatgatgc agaattaaca 240
ttaaataaag tggatgaaga atttgtggaa gaagagacag ataatgaaga aaactttatt 300
gatctcaacg ttttaaaggc ccagacatat cacttggata tgaacgagac tgccaaacaa 360
gaagatattt tggaatccac aacagatgct gcagaatgga gcctagaagt ggaacgtgta 420
ctaccgcaac tgaaagtcac gattaggact gacaataagg attggagaat ccatgttgac 480
caaatgcacc agcacagaag tggaattgaa tctgctctaa aggagaccaa gggatttttg 540
gacaaactcc ataatgaaat tactaggact ttggaaaaga tcagcagccg agaaaagtac 600
atcaacaatc agccgggagc ccatggagca ctgtcctcag agatgcgcag gttaggctca 660
ctgtctaggc caggcccacc ttagtcactg tggactggca atggaagctc ttcctggaca 720
cacctgccct agccctcacc ctggggtgga agagaaatga gcttggcttg caactcagac 780
cattccacgg aggcatcctc cccttccctg ggctggtgaa taaaagtttc ctgaggtcaa 840
ggacttcctt ttccctgcca aaatggtgtc cagaactttg aggccagagg tgatccagtg 900
atttgggagc tgcaggtcac acaggctgct cagagggctg ctgaacagga tgtcctcgga 960
cgacaggcac ctgggctcca gctgcggctc cttcatcaag actgagccgt ccagcccgtc 1020
ctcgggcata gatgccctca gccaccacag ccccagtggc tcgtccgacg ccagcggcgg 1080
ctttggcctg gccctgggca cccacgccaa cggtctggac tcgccaccca tgtttgcagg 1140
cgccgggctg ggaggcaccc catgccgcaa gagctacgag gactgtgcca gcggcatcat 1200

RAW SEQUENCE LISTING DATE: 04/25/2002
PATENT APPLICATION: US/10/054,841 TIME: 12:04:27

Input Set : A:\19999YCA.txt
Output Set: N:\CRF3\04252002\J054841.raw

```
59 ggaggactcg gccatcaagt gcgagtacat gctcaacgcc atccccaagc gcctgtgcct    1260
60 cgtgtgcggg gacattgcct ctggctacca ctacggcgtg gcctcctgcg aggcttgcaa    1320
61 ggccttcttc aagaggacta tccaagggaa cattgagtac agctgcccgg ccaccaacga    1380
62 gtgcgagatc accaaacgga ggcgcaagtc ctgccaggcc tgccgcttca tgaaatgcct    1440
63 caaagtgggg atgctgaagg aaggtgtgcg ccttgatcga gtgcgtggag gccgtcagaa    1500
64 atacaagcga cggctggact cagagagcag cccatacctg agcttacaaa tttctccacc    1560
65 tgctaaaaag ccattgacca agattgtctc atacctactg gtggctgagc cggacaagct    1620
66 ctatgccatg cctccccctg gtatgcctga gggggacatc aaggccctga ccactctctg    1680
67 tgacctggca gaccgagagc ttgtggtcat cattggctgg gccaagcaca tcccaggctt    1740
68 ctcaagcctc tccctggggg accagatgag cctgctgcag agtgcctgga tggaaatcct    1800
69 catcctgggc atcgtgtacc gctcgctgcc ctacgacgac aagctggtgt acgctgagga    1860
70 ctacatcatg gatgaggagc actcccgcct cgcggggctg ctggagctct accgggccat    1920
71 cctgcagctg gtacgcaggt acaagaagct caaggtggag aaggaggagt ttgtgacgct    1980
72 caaggccctg gccctcgcca actccgattc catgtacatc gaggatctag aggctgtcca    2040
73 gaagctgcag gacctgctgc acgaggcact gcaggactac gagctgagcc agcgccatga    2100
74 ggagccctgg aggacgggca agctgctgct gacactgccg ctgctgcggc agacggccgc    2160
75 caaggccgtg cagcacttct atagcgtcaa actgcagggc aaagtgccca tgcacaaact    2220
76 cttcctggag atgctggagg ccaaggcctg ggccagggct gactcccttc aggagtggag    2280
77 gccactggag caagtgccct ctcccctcca ccgagccacc aagaggcagc atgtgcattt    2340
78 cctaactccc ttgccccctc ccccatctgt ggcctgggtg ggcactgctc aggctggata    2400
79 ccacctggag gttttccttc cgcagagggc aggttggcca agagcagctt agaggatctc    2460
80 ccaaggatga aagaatgtca agccatgatg gaaaatgccc cttccaatca gctgccttca    2520
81 caagcaggga tcagagcaac tccccgggga tccccaatcc acgcccttct agtccaaccc    2580
82 ccctcaatga gagaggcagg cagatctcac ccagcactag gacaccagga ggccagggaa    2640
83 agcatctctg gctcaccatg taacatctgg cttggagcaa gtgggtgttc tgcacaccag    2700
84 gcagctgcac ctcactggat ctagtgttgc tgcgagtgac ctcacttcag agcccctcta    2760
85 gcagagtggg gcggaagtcc tgatggttgg tgtccatgag gtggaag                  2807
87 <210> SEQ ID NO: 2
88 <211> LENGTH: 500
89 <212> TYPE: PRT
90 <213> ORGANISM: Human
92 <400> SEQUENCE: 2
93 Met Ser Ser Asp Asp Arg His Leu Gly Ser Ser Cys Gly Ser Phe Ile
94  1               5                  10                  15
95 Lys Thr Glu Pro Ser Ser Pro Ser Ser Gly Ile Asp Ala Leu Ser His
96             20                  25                  30
97 His Ser Pro Ser Gly Ser Ser Asp Ala Ser Gly Gly Phe Gly Leu Ala
98         35                  40                  45
99 Leu Gly Thr His Ala Asn Gly Leu Asp Ser Pro Pro Met Phe Ala Gly
100     50                  55                  60
101 Ala Gly Leu Gly Gly Thr Pro Cys Arg Lys Ser Tyr Glu Asp Cys Ala
102 65                  70                  75                  80
103 Ser Gly Ile Met Glu Asp Ser Ala Ile Lys Cys Glu Tyr Met Leu Asn
104                 85                  90                  95
105 Ala Ile Pro Lys Arg Leu Cys Leu Val Cys Gly Asp Ile Ala Ser Gly
106             100                 105                 110
107 Tyr His Tyr Gly Val Ala Ser Cys Glu Ala Cys Lys Ala Phe Phe Lys
108         115                 120                 125
109 Arg Thr Ile Gln Gly Asn Ile Glu Tyr Ser Cys Pro Ala Thr Asn Glu
```

RAW SEQUENCE LISTING DATE: 04/25/2002
PATENT APPLICATION: US/10/054,841 TIME: 12:04:27

Input Set : A:\19999YCA.txt
Output Set: N:\CRF3\04252002\J054841.raw

```
110      130                 135                 140
111 Cys Glu Ile Thr Lys Arg Arg Arg Lys Ser Cys Gln Ala Cys Arg Phe
112 145                 150                 155                 160
113 Met Lys Cys Leu Lys Val Gly Met Leu Lys Glu Gly Val Arg Leu Asp
114                 165                 170                 175
115 Arg Val Arg Gly Gly Arg Gln Lys Tyr Lys Arg Arg Leu Asp Ser Glu
116             180                 185                 190
117 Ser Ser Pro Tyr Leu Ser Leu Gln Ile Ser Pro Pro Ala Lys Lys Pro
118         195                 200                 205
119 Leu Thr Lys Ile Val Ser Tyr Leu Leu Val Ala Glu Pro Asp Lys Leu
120     210                 215                 220
121 Tyr Ala Met Pro Pro Pro Gly Met Pro Glu Gly Asp Ile Lys Ala Leu
122 225                 230                 235                 240
123 Thr Thr Leu Cys Asp Leu Ala Asp Arg Glu Leu Val Val Ile Ile Gly
124                 245                 250                 255
125 Trp Ala Lys His Ile Pro Gly Phe Ser Ser Leu Ser Leu Gly Asp Gln
126             260                 265                 270
127 Met Ser Leu Leu Gln Ser Ala Trp Met Glu Ile Leu Ile Leu Gly Ile
128         275                 280                 285
129 Val Tyr Arg Ser Leu Pro Tyr Asp Asp Lys Leu Val Tyr Ala Glu Asp
130     290                 295                 300
131 Tyr Ile Met Asp Glu Glu His Ser Arg Leu Ala Gly Leu Leu Glu Leu
132 305                 310                 315                 320
133 Tyr Arg Ala Ile Leu Gln Leu Val Arg Arg Tyr Lys Lys Leu Lys Val
134                 325                 330                 335
135 Glu Lys Glu Glu Phe Val Thr Leu Lys Ala Leu Ala Leu Ala Asn Ser
136             340                 345                 350
137 Asp Ser Met Tyr Ile Glu Asp Leu Glu Ala Val Gln Lys Leu Gln Asp
138         355                 360                 365
139 Leu Leu His Glu Ala Leu Gln Asp Tyr Glu Leu Ser Gln Arg His Glu
140     370                 375                 380
141 Glu Pro Trp Arg Thr Gly Lys Leu Leu Leu Thr Leu Pro Leu Leu Arg
142 385                 390                 395                 400
143 Gln Thr Ala Ala Lys Ala Val Gln His Phe Tyr Ser Val Lys Leu Gln
144                 405                 410                 415
145 Gly Lys Val Pro Met His Lys Leu Phe Leu Glu Met Leu Glu Ala Lys
146             420                 425                 430
147 Ala Trp Ala Arg Ala Asp Ser Leu Gln Glu Trp Arg Pro Leu Glu Gln
148         435                 440                 445
149 Val Pro Ser Pro Leu His Arg Ala Thr Lys Arg Gln His Val His Phe
150     450                 455                 460
151 Leu Thr Pro Leu Pro Pro Pro Pro Ser Val Ala Trp Val Gly Thr Ala
152 465                 470                 475                 480
153 Gln Ala Gly Tyr His Leu Glu Val Phe Leu Pro Gln Arg Ala Gly Trp
154                 485                 490                 495
155 Pro Arg Ala Ala
156             500
158 <210> SEQ ID NO: 3
159 <211> LENGTH: 2985
```

RAW SEQUENCE LISTING DATE: 04/25/2002
PATENT APPLICATION: US/10/054,841 TIME: 12:04:27

Input Set : A:\19999YCA.txt
Output Set: N:\CRF3\04252002\J054841.raw

<212> TYPE: DNA
<213> ORGANISM: Human
<400> SEQUENCE: 3

gcgggccgcc agtgtggtgg aattcggctt gtcactagga gaacatttgt gttaattgca 60
ctgtgctctg tcaaggaaac tttgatttat agctggggtg cacaaataat ggttgccggt 120
cgcacatgga ttcggtagaa ctttgccttc ctgaatcttt ttccctgcac tacgaggaag 180
agcttctctg cagaatgtca aacaaagatc gacacattga ttccagctgt tcgtccttca 240
tcaagacgga accttccagc ccagcctccc tgacggacag cgtcaaccac cacagccctg 300
gtggctcttc agacgccagt gggagctaca gttcaaccat gaatggccat cagaacggac 360
ttgactcgcc acctctctac ccttctgctc ctatcctggg aggtagtggg cctgtcagga 420
aactgtatga tgactgctcc agcaccattg ttgaagatcc ccagaccaag tgtgaataca 480
tgctcaactc gatgcccaag agactgtgtt tagtgtgtgg tgacatcgct tctgggtacc 540
actatggggt agcatcatgt gaagcctgca aggcattctt caagaggaca attcaaggca 600
atatagaata cagctgccct gccacgaatg aatgtgaaat cacaaagcgc agacgtaaat 660
cctgccaggc ttgccgcttc atgaagtgtt taaaagtggg catgctgaaa gaaggggtgc 720
gtcttgacag agtacgtgga ggtcggcaga agtacaagcg caggatagat gcggagaaca 780
gcccatacct gaaccctcag ctggttcagc cagccaaaaa gccatataac aagattgtct 840
cacatttgtt ggtggctgaa ccggagaaga tctatgccat gcctgaccct actgtccccg 900
acagtgacat caaagccctc actacactgt gtgacttggc cgaccgagag ttggtggtta 960
tcattggatg ggcgaagcat attccaggct tctccacgct gtccctggcg gaccagatga 1020
gccttctgca gagtgcttgg atggaaattt tgatccttgg tgtcgtatac cggtctcttt 1080
catttgagga tgaacttgtc tatgcagacg attatataat ggacgaagac cagtccaaat 1140
tagcaggcct tcttgatcta aataatgcta tcctgcagct ggtaaagaaa tacaagagca 1200
tgaagctgga aaaagaagaa tttgtcaccc tcaaagctat agctcttgct aattcagact 1260
ccatgcacat agaagatgtt gaagccgttc agaagcttca ggatgtctta catgaagcgc 1320
tgcaggatta tgaagctggc cagcacatgg aagaccctcg tcgagctggc aagatgctga 1380
tgacactgcc actcctgagg cagacctcta ccaaggccgt gcagcatttc tacaacatca 1440
aactagaagg caaagtccca atgcacaaac ttttttgga aatgttggag gccaaggtct 1500
gactaaaagc tccctgggcc ttcccatcct tcatgttgaa aaagggaaaa taaacccaag 1560
agtgatgtcg aagaaactta gagtttagtt aacaacatca aaaatcaaca gactgcactg 1620
ataatttagc agcaagacta tgaagcagct ttcagattcc tccataggtt cctgatgagt 1680
tctttctact ttctccatca tcttctttcc tctttcttcc cacatttctc tttctcttta 1740
ttttttctcc ttttcttctt tcacctccct tatttctttg cttctttcat tcctagttcc 1800
cattctcctt tattttcttc ccgtctgcct gccttctttc ttttctttac ctactctcat 1860
tcctctcttt tctcatcctt ccccttttct ctaaatttga aatagcttta gtttaaaaaa 1920
aaaaatcctc ccttcccccc ttcctttccc tttctttcct ttttcccttt ccttttccct 1980
ttcctttcct ttcctcttga ccttctttcc atctttcttt ttcttccttc tgctgctgaa 2040
cttttaaaag aggtctctaa ctgaagagag atggaagcca gccctgccaa aggatggaga 2100
tccataatat ggatgccagt gaacttattg tgaaccatac cgtccccaat gactaaggaa 2160
tcaaagagag agaaccaacg ttcctaaaag tacagtgcaa catatacaaa ttgactgagt 2220
gcagtattag atttcatggg agcagcctct aattagacaa cttaagcaac gttgcatcgg 2280
ctgcttctta tcattgcttt tccatctaga tcagttacag ccatttgatt ccttaattgt 2340
cttttcaagt cttccaggta tttgttagtt tagctactat gtaacttttt cagggaatag 2400
tttaagcttt attcattcat gcaatactaa agagaaataa gaatactgca attttgtgct 2460
ggctttgaac aattacgaac aataatgaag gacaaatgaa tcctgaagga agatttttaa 2520
aaatgttttg tttcttctta caaatggaga tttttttgta ccagctttac cacttttcag 2580
ccatttatta atatgggaat ttaacttact caagcaatag ttgaagggaa ggtgcatatt 2640
atcacggatg caatttatgt tgtgtgccag tctggtccca aacatcaatt tcttaacatg 2700
agctccagtt tacctaaatg ttcactgaca caaaggatga gattacacct acagtgactc 2760

RAW SEQUENCE LISTING DATE: 04/25/2002
PATENT APPLICATION: US/10/054,841 TIME: 12:04:27

Input Set : A:\19999YCA.txt
Output Set: N:\CRF3\04252002\J054841.raw

```
tgagtagtca catatataag cactgcacat gagatataga tccgtagaat tgtcaggagt    2820
gcacctctct acttgggagg tacaattgcc atatgatttc tagctgccat ggtggttagg    2880
aatgtgatac tgcctgtttg caaagtcaca gaccttgcct cagaaggagc tgtgagccag    2940
tattcattta agagaattcc accacactgg cggcccgcgc ttgat                    2985
<210> SEQ ID NO: 4
<211> LENGTH: 458
<212> TYPE: PRT
<213> ORGANISM: Human
<400> SEQUENCE: 4
Met Asp Ser Val Glu Leu Cys Leu Pro Glu Ser Phe Ser Leu His Tyr
 1               5                  10                  15
Glu Glu Glu Leu Leu Cys Arg Met Ser Asn Lys Asp Arg His Ile Asp
            20                  25                  30
Ser Ser Cys Ser Ser Phe Ile Lys Thr Glu Pro Ser Ser Pro Ala Ser
        35                  40                  45
Leu Thr Asp Ser Val Asn His His Ser Pro Gly Gly Ser Ser Asp Ala
    50                  55                  60
Ser Gly Ser Tyr Ser Ser Thr Met Asn Gly His Gln Asn Gly Leu Asp
65                  70                  75                  80
Ser Pro Pro Leu Tyr Pro Ser Ala Pro Ile Leu Gly Gly Ser Gly Pro
                85                  90                  95
Val Arg Lys Leu Tyr Asp Asp Cys Ser Ser Thr Ile Val Glu Asp Pro
            100                 105                 110
Gln Thr Lys Cys Glu Tyr Met Leu Asn Ser Met Pro Lys Arg Leu Cys
        115                 120                 125
Leu Val Cys Gly Asp Ile Ala Ser Gly Tyr His Tyr Gly Val Ala Ser
    130                 135                 140
Cys Glu Ala Cys Lys Ala Phe Phe Lys Arg Thr Ile Gln Gly Asn Ile
145                 150                 155                 160
Glu Tyr Ser Cys Pro Ala Thr Asn Glu Cys Glu Ile Thr Lys Arg Arg
                165                 170                 175
Arg Lys Ser Cys Gln Ala Cys Arg Phe Met Lys Cys Leu Lys Val Gly
            180                 185                 190
Met Leu Lys Glu Gly Val Arg Leu Asp Arg Val Arg Gly Gly Arg Gln
        195                 200                 205
Lys Tyr Lys Arg Arg Ile Asp Ala Glu Asn Ser Pro Tyr Leu Asn Pro
    210                 215                 220
Gln Leu Val Gln Pro Ala Lys Lys Pro Tyr Asn Lys Ile Val Ser His
225                 230                 235                 240
Leu Leu Val Ala Glu Pro Glu Lys Ile Tyr Ala Met Pro Asp Pro Thr
                245                 250                 255
Val Pro Asp Ser Asp Ile Lys Ala Leu Thr Thr Leu Cys Asp Leu Ala
            260                 265                 270
Asp Arg Glu Leu Val Val Ile Ile Gly Trp Ala Lys His Ile Pro Gly
        275                 280                 285
Phe Ser Thr Leu Ser Leu Ala Asp Gln Met Ser Leu Leu Gln Ser Ala
    290                 295                 300
Trp Met Glu Ile Leu Ile Leu Gly Val Val Tyr Arg Ser Leu Ser Phe
305                 310                 315                 320
```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 04/25/2002
PATENT APPLICATION: US/10/054,841 TIME: 12:04:28

Input Set : A:\19999YCA.txt
Output Set: N:\CRF3\04252002\J054841.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 279,294
Seq#:8; N Pos. 2,3,4,5,6,7,8,9,10,11,12,13,18,19,37,42,401,471,521,532,534
Seq#:8; N Pos. 550,556,557,558,559,560,561,562,563,564,565,566,567,568,569
Seq#:8; N Pos. 570,571,572,573,574,575,576,577,578,579,580,581,582,583,584
Seq#:8; N Pos. 585,586,587,588,589,590,591,592,593,594,595,596,597,598,599
Seq#:8; N Pos. 600,601,602,603,604,605,606,607,608,609,610,611,612,613,614
Seq#:8; N Pos. 615,616,617,618,619,620

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/054,841

DATE: 04/25/2002
TIME: 12:04:28

Input Set : A:\19999YCA.txt
Output Set: N:\CRF3\04252002\J054841.raw

L:413 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:240
L:428 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0
L:434 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:360
L:435 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:420
L:436 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:480
L:437 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:540
L:438 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:600